**Improving Language Understanding by Generative**

**Pre-Training**

**Alec Radford**

OpenAI alec@openai.com

**Karthik Narasimhan**

OpenAI karthikn@openai.com

**Tim Salimans**

OpenAI tim@openai.com

**Ilya Sutskever**

OpenAI ilyasu@openai.com

**Abstract**

Natural language understanding comprises a wide range of diverse tasks such as textual entailment, question answering, semantic similarity assessment, and document classification. Although large unlabeled text corpora are abundant, labeled data for learning these specific tasks is scarce, making it challenging for discriminatively trained models to perform adequately. We demonstrate that large gains on these tasks can be realized by generative pre-training of a language model on a diverse corpus of unlabeled text, followed by discriminative fine-tuning on each specific task. In contrast to previous approaches, we make use of task-aware input transformations during fine-tuning to achieve effective transfer while requiring minimal changes to the model architecture. We demonstrate the effectiveness of our approach on a wide range of benchmarks for natural language understanding. Our general task-agnostic model outperforms discriminatively trained models that use architectures specifically crafted for each task, significantly improving upon the state of the art in 9 out of the 12 tasks studied. For instance, we achieve absolute improvements of 8.9% on commonsense reasoning (Stories Cloze Test), 5.7% on question answering (RACE), and 1.5% on textual entailment (MultiNLI).

**摘要**

自然语言理解包括各种各样的任务，例如文本蕴涵，问答，语义相似性评估和文档分类等。虽然大量未标记的文本语料库很丰富，但用于学习这些特定任务的标记数据很少，这使得有条件训练的模型难以充分发挥作用。我们证明，通过对多种未标记文本语料库的语言模型进行生成预训练，然后对每项特定任务进行辨别性微调，可以实现这些任务的巨大收益。与以前的方法相比，我们在微调期间利用任务感知输入转换来实现有效传输，同时对模型架构进行最少的更改。我们证明了我们的方法在广泛的自然语言理解基准上的有效性。我们的一般任务不可知模型优于使用专门为每项任务设计的架构的经过训练的训练模型，在所研究的12项任务中的9项中显着改进了现有技术水平。例如，我们在常识推理（Stories Cloze Test）上获得8.9％的绝对改善，在问答（RACE）上达到5.7％，在文本蕴涵（MultiNLI）上达到1.5％。